Uniform Convergence of Semimartingales and Minimum Distance Estimators

C. Sibeux. Laboratoire Statistique et Processus, Département de Mathématiques, Université d'Angers, 2 Boulevard Lavoisier, 49045 Angers cedex, FRANCE
L. Vostrikova. Laboratoire Statistique et Processus, Département de Mathématiques, Université d'Angers, 2 Boulevard Lavoisier, 49045 Angers cedex, FRANCE

Abstract

We consider a sequence of special semimartingales $X^{n,\theta} = (X_t^{n,\theta})_{t\geq 0}$ depending on the parameter $\theta \in \Theta \subseteq I\!R^m$ and give conditions, expressed in terms of the predictable characteristics of $X^{n,\theta}$, for the uniform in θ, weak convergence of semimartingales $X^{n,\theta}$. Using results about uniform convergence we study the minimal distance estimators of θ for semimartingales.

Key words : uniform weak convergence, semimartingale, predictable characteristic, minimum distance estimator, consistency, asymptotic normality.

AMS classification : 60F17, 62F12 .

1 Introduction

Uniform (with respect to the parameter) weak convergence of stochastic processes plays an important role in statistics, for instance in studying of the weak convergence of the maximum likelihood estimators, bayesian estimators, minimum distance estimators, for obtaining the minimax theorems for estimators etc. We mention the paper of Jacod, Mano [9] where the uniform convergence of the processes was equivalent to the uniform convergence of the Prohorov distances between them, the book of Shiryayev, Spokoinij [24] where the notion of the λ-convergence was introduced, useful for the minimax theorems of estimators, as well as the papers of Vostrikova [26], [27] where a weak convergence of the likelihood ratio processes was studied in the Skorohod space $D(\mathbb{R}^+, C_{loc})$ with the values in the space of the continuous functions endowed with the locally-uniform metric.

We can often suppose that we observe the sequence of special semimartingales $X^{n,\theta} = (X_t^{n,\theta})_{t\geq 0}$ depending on the parameter $\theta \in \Theta \subseteq \mathbb{R}^m$. For instance, if we consider the empirical process for the first [nt] observations of a sequence of independent random variables $(X_k)_{k\geq 1}$ having the same distribution function F_θ, then the process $X^{n,\theta}$ will be given by

$$X_t^{n,\theta}(u) = \frac{1}{n}\sum_{k=1}^{[nt]} I_{\{X_k \leq u\}}$$

where $t \in [0,1]$, $u \in \mathbb{R}$ and $I_{\{.\}}$ is an indicator function. The process $X^{n,\theta}(u) = (X_t^{n,\theta}(u))_{0\leq t\leq 1}$ is a special semimartingale on the naturally defined probability filtered space, and its canonical decomposition is given by:

$$X_t^{n,\theta}(u) = \frac{1}{n}\sum_{k=1}^{[nt]}(I_{\{X_k \leq u\}} - F_\theta(u)) + \frac{[nt]}{n}F_\theta(u)$$

If we consider a sequence of the diffusion type processes with small noise, then under some natural conditions they will be special semimartingales also.

One of the aims of this paper is to give the conditions for a weak convergence in $D(\mathbb{R}^+, C_{loc})$ of a sequence of the special semimartingales $X^{n,\theta}$ to a continuous gaussian semimartingale $X^\theta = (X_t^\theta)_{t\geq 0}$. Here we mention the paper of C. Sibeux [25], where such conditions were obtained in the case of the diffusion type limit process X^θ.

But what is the relation between the convergence in $D(\mathbb{R}^+, C_{loc})$ and the convergence of the estimators of type arg sup or arg inf like maximum likelihood estimators or minimum distance estimators? Using the Skorohod representation theorem one can see that the convergence in $D(\mathbb{R}^+, C_{loc})$ implies uniform on compact sets convergence of processes which, under some supplementary conditions (as the ones of lemma 3 of 4.) assuring in fact the continuity of arg sup or arg inf, gives the convergence of the estimators.

Minimum distance estimators were studied by many authors. Among the papers published on this subjet we mention the papers of Wolfowitz [28], La Riccia [20], Bolthausen [3], Koul and Wet[11], Millar[22][23]; some other references we can find in Parr[18]. Notably we draw attention to the paper of Millar[23] "A General Approach of the Optimality of Minimum Distance Estimators", where one can find the general conditions on the observed processes providing the consistency, the asymptotic normality and the optimality of the minimum distance estimators. For stochastic processes the minimum distance estimators were studied: for diffusion processes in the paper of Kutoyants [12], Kutoyants and Pilibossian [17], Henaff [6], for the diffusion processes with delay in the papers of Kutoyants and Mourid [16], for the diffusion ergodic processes in Fournie, Kutoyants [5], for spacial Poisson processes by Kutoyants, Liese [15], for partially observed linear stochastic systems by Bertrand and Kutoyants [1], for recursive markov processes by Höpfner, Kutoyants [7], for diffusion random fields by Kutoyants, Lessi [14]. Most of these results can also be found in the book of Kutoyants [13].

From the examples given in the paper of Millar [23] and in the book of Kutoyants [13], we can deduce that a natural formulation of the minimum distance problem for semimartingales could be the following. We suppose that on the probability space $(\Omega, \mathcal{F}, I\!\!F, P^{\epsilon})$ with the filtration $I\!\!F = (\mathcal{F}_t)_{t\geq 0}$ we have the sequence of the special semimartingales $X^{\epsilon,\theta} = (X_t^{\epsilon,\theta})_{t\geq 0}$ such that

$$X_t^{\epsilon,\theta} = X_0^{\epsilon,\theta} + A_t^{\epsilon,\theta} + M_t^{\epsilon,\theta} \tag{1}$$

where $A^{\epsilon,\theta} = (A_t^{\epsilon,\theta})_{t\geq 0}$ is a predictable process of the locally integrable variation and $M^{\epsilon,\theta} = (M_t^{\epsilon,\theta})_{t\geq 0}$ is a local martingale. We suppose that as $\epsilon \to 0$ the $X^{\epsilon,\theta}$ converge weakly of some deterministic cadlag function of locally bounded variation $X^{\theta} = (X_t^{\theta})_{t\geq 0}$, so that $X^{\epsilon,\theta}$ can be interpreted as the observation of X^{θ} with the systematic error $A^{\epsilon,\theta} - X^{\theta}$ and the martingale noise $M^{\epsilon,\theta}$.

We suppose that in the space of the trajectories $D([0,T], I\!\!R)$ with $T > 0$, we fixed the norm $|| \quad ||_T$. Then we can define the minimum distance estimator as follows:

$$\widehat{\theta}_T^{\epsilon} = \arg \inf_{\theta\in\Theta} || X^{\epsilon,\theta_0} - X^{\theta} ||_T$$

with the following convention about $arg \inf$: if inf is not achieved it is equal to ∞ and if it is achieved in several points we choose one arbitrarily. The aim of this paper is to find conditions expressed in terms of the triplets of the observed semimartingales, under which the minimum distance estimators are consistent and asymptotically normal.

In 2. we give the criteria for convergence in the space $D(I\!\!R^+, C_{loc})$. In 3. we express the conditions in terms of the triplets of semimartingales providing a weak convergence in $D(I\!\!R^+, C_{loc})$ (see theorem 3). In 4. we give the conditions which imply consistency (see theorem 4) and a weak convergence of the normalized estimators (see theorem 5). For the L_2-norm we obtain under the conditions of the theorem 5 the asymptotic normality of the minimum distance estimators (see corollary 1).

The main tool used in this paper is the general theory of stochastic processes, and we refer the readers for the technical details to the books of Jacod [8], Jacod, Shiryayev [10], Liptser, Shiryayev [21].

2 The Space $D(I\!\!R^+, C_{loc})$

We consider the Skorohod space $D(I\!\!R^+, C_{loc})$ with values in the space of continuous functions $C_{loc}(I\!\!R^m)$ endowed with the locally uniform metric. We recall that if $(K_i)_{i\geq 1}$ is a sequence of increasing compact sets from $I\!\!R^m$, then the distance d(X,Y) between two elements $X = (X^\theta)_{\theta\in\mathrm{R}^m}$ and $Y = (Y^\theta)_{\theta\in\mathrm{R}^m}$ of the space $C_{loc}(I\!\!R^m)$ can be defined by:

$$d(X,Y) = \sum_{i=1}^{+\infty} 2^{-i} \frac{\sup\limits_{\theta\in K_i} \mid X^\theta - Y^\theta \mid}{1 + \sup\limits_{\theta\in K_i} \mid X^\theta - Y^\theta \mid}.$$

It is well known that the space $D(I\!\!R^+, C_{loc})$ endowed with the Skorohod distance is a complete separable space, since $C_{loc}(I\!\!R^m)$ is itself complete and separable . This fact permits us to use Prohorov's theorem [2] about tightness and relative compactness to obtain criteria for a weak convergence.

To characterize the compact sets of the space $D(I\!\!R^+, C_{loc})$, we introduce the following moduli of continuity the first of which is the standard modulus of the continuity in the space $D(I\!\!R^+, I\!\!R)$ applied to X^θ with a fixed value of $\theta \in I\!\!R^m$.

Denoting by $\mathcal{G}$ the set of the partitions $\{0 = t_0 < t_1 < \ldots < t_n < N\}$ of the interval [0,N] with $N > 0$, satisfying the condition $\min\limits_{0\leq j\leq n-1}(t_{j+1} - t_j) > h$, we set

$$W_h^N(X^\theta) = \inf_{\mathcal{G}} \max_{0 \le j \le n} \sup_{s,t \in [t_j, t_{j+1}[} \mid X_s^\theta - X_t^\theta \mid$$

and let

$$V_h^{N,i}(X) = \sup_{0 \le t \le N} \sup_{\substack{\theta,\theta' \in K_i \\ |\theta - \theta'| \le h}} \mid X_t^\theta - X_t^{\theta'} \mid .$$

where K_i is a compact set for the above definition of the distance in the space $C_{loc}(I\!R^m)$.

The following theorem gives the necessary and sufficient conditions for the relative compactness in $D(I\!R^+, C_{loc})$.

Theorem 1 *(see [26],[25])A subset $\mathcal{K}$ of $D(I\!R^+, C_{loc})$ is relatively compact if and only if the following conditions are satisfied: for each $i \ge 1$ and $N > 0$*

1) $\sup_{X \in \mathcal{K}} \sup_{0 \le t \le N} \sup_{\theta \in K_i} \mid X_t^\theta \mid < \infty$,

2) $\lim_{h \to 0} \sup_{X \in \mathcal{K}} \sup_{\theta \in K_i} W_h^N(X^\theta) = 0$,

3) $\lim_{h \to 0} \sup_{X \in \mathcal{K}} V_h^{N,i}(X) = 0$.

Now suppose that we have a sequence of stochastic processes $X^n = (X_t^{n,\theta})_{t \ge 0, \theta \in \mathrm{R}^m}$ and a "limit" process $X = (X_t^\theta)_{t \ge 0, \theta \in \mathrm{R}^m}$ with trajectories in the space $D(I\!R^+, C_{loc})$, given on the probability spaces $(\Omega^n, \mathcal{F}^n, P^n)$ and $(\Omega, \mathcal{F}, P)$ respectively. Then using Prohorov's theorem [2] about the tightness and relative compactness and theorem 1 we can obtain the following.

Theorem 2 *(see [26],[25]). We assume that finite-dimensional distributions of X^n converge weakly to finite-dimensional distributions of X. We suppose also that for each $\epsilon > 0$ and $N > 0$*

1) $\lim_{h \to 0} \overline{lim}_{n \to +\infty} P^n(W_h^N(X^{n,\theta}) \ge \epsilon) = 0, \forall \theta \in I\!R^m$,

2) $\lim_{h \to 0} \sup_{n \ge 1} P^n(V_h^{N,i}(X^n) \ge \epsilon) = 0$.

Then we have a weak convergence of X^n to of X in the Skorohod space $D(I\!R^+, C_{loc})$:

$$X^n \stackrel{w(P^n)}{\longrightarrow} X.$$

3 Weak Convergence of Semimartingales in the space $D(I\!R^+, C_{loc})$.

We suppose that we are given probability spaces $(\Omega^n, \mathcal{F}^n, I\!F^n, P^n)$ and $(\Omega, \mathcal{F}, I\!F, P)$ endowed with the filtrations $I\!F^n = (\mathcal{F}_t^n)_{t\geq 0}$ and $I\!F = (\mathcal{F}_t)_{t\geq 0}$. Let $X^{n,\theta} = (X_t^{n,\theta})_{t\geq 0}$ and $X^\theta = (X_t^\theta)_{t\geq 0}$ be semimartingales depending on the parameter $\theta \in I\!R^m$, defined on the spaces $(\Omega^n, \mathcal{F}^n, I\!F^n, P^n)$ and $(\Omega, \mathcal{F}, I\!F, P)$ respectively.

We suppose that the semimartingales $X^{n,\theta}$ are special and locally square-integrable, and that the semimartingale X^θ is continuous with a deterministic initial value. This means that we have a unique decomposition:

$$X_t^{n,\theta} = X_0^{n,\theta} + M_t^{n,\theta} + A_t^{n,\theta}, \;\; X_t^\theta = X_0^\theta + M_t^\theta + A_t^\theta$$

where $M^{n,\theta} = (M_t^{n,\theta})_{t\geq 0}$ is a locally square-integrable martingale, $M^\theta = (M_t^\theta)_{t\geq 0}$ is a continuous martingale, and $A^{n,\theta} = (A_t^{n,\theta})_{t\geq 0}$ and $A^\theta = (A_t^\theta)_{t\geq 0}$ are predictable processes with locally integrable variation.

If we denote by $\mu^{n,\theta}$ the jump's measure of the semimartingale $X^{n,\theta}$ and by $\nu^{n,\theta}$ its compensator, then we can write the following decomposition (see Jacod [8]):

$$X_t^{n,\theta} = X_0^{n,\theta} + A_t^{n,\theta} + X_t^{n,\theta,c} + \int_0^t \int\limits_{\mathrm{R}\setminus\{0\}} xd(\mu^{n,\theta} - \nu^{n,\theta}),$$

$$X_t^\theta = X_0^\theta + A_t^\theta + X_t^{\theta,c},$$

where $X^{n,\theta,c} = (X_t^{n,\theta,c})_{t\geq 0}$ and $X^{\theta,c} = (X_t^{\theta,c})_{t\geq 0}$ are continuous martingale components of $X^{n,\theta}$ and X^θ respectively.

We denote by $< X^{n,\theta,c} >$ and $< X^{\theta,c} >$ the predictable variation of the corresponding processes.

The aim of this part is to give conditions expressed in terms of the predictable characteristics $(A^{n,\theta}, < X^{n,\theta,c} >, \nu^{n,\theta})$ and $(A^\theta, < X^{\theta,c} >, \nu^\theta)$ of the semimartingales $X^{n,\theta}$ and X^θ respectively, providing the existence of modifications of the processes $X^n = (X_t^{n,\theta})_{t\geq 0, \theta\in \mathrm{R}^m}$, $X = (X_t^\theta)_{t\geq 0,\theta\in\mathrm{R}^m}$ with trajectories in $D(I\!R^+, C_{loc})$ for which a weak convergence in $D(I\!R^+, C_{loc})$ takes place.

We suppose that the standard conditions on the triplets of the "limit" semimartingale, given in the book of Liptser, Shiryayev [21], are satisfied.

Namely, we suppose that for each $\theta \in I\!R^m$, the predictable characteristics as well as their linear combinations, determine uniquely the solution of the corresponding semimartingale problem and that for all $t \geq 0$

$$A_t^\theta(X) = \int_0^t a^\theta(s,X)du_s, \tag{2}$$

$$< M^\theta >_t (X) = \int_0^t c^\theta(s,X)du_s, \tag{3}$$

$$< M^\theta, M^{\theta'} >_t (X,Y) = \int_0^t d^{\theta,\theta'}(s,X,Y)du_s. \tag{4}$$

We assume that the functions $a^\theta(t,X)$, $c^\theta(t,X)$, $d^{\theta,\theta'}(t,X,Y)$ are predictable and that their moduli are integrable with respect to du_s where $(u_s)_{s\geq 0}$ is some continuous increasing deterministic function of locally bounded variation.

We suppose that the functions $a^\theta(t,X)$, $c^\theta(t,X)$ and $d^{\theta,\theta'}(t,X,Y)$ are continuous in the Skorohod metric and that they verify

$$| a^\theta(t,X) | \leq L(t)(1 + \sup_{s<t} | X_s |), \tag{5}$$

$$| c^\theta(t,X) | \leq L(t)(1 + \sup_{s<t} | X_s |^2), \tag{6}$$

$$| d^{\theta,\theta'}(t,X,Y) | \leq L(t)(1 + \sup_{s<t} | X_s | . \sup_{s<t} | Y_s |), \tag{7}$$

where L(t) is positive and integrable on the finite intervals function which can depend on θ and θ'.

We suppose further that the following conditions are satisfied:

Conditions of group 1. For every $\theta \in I\!R^m$, $\epsilon > 0$ and $T > 0$ we have:

1) $X_0^{n,\theta} \xrightarrow{P^n} X_0^\theta$,

2) $\overline{\lim}_{n\to+\infty} P^n(\int_0^T \int_{|x|>\delta} x^2 d\nu^{n,\theta} \geq \epsilon) = 0, \forall \delta > 0,$

3) $\overline{\lim\limits_{n\to+\infty}} P^n(\sup\limits_{t\leq T} \mid A_t^{n,\theta} - \int\limits_0^t a^\theta(s, X^{n,\theta})du_s \mid\geq \epsilon) = 0,$

4) $\overline{\lim\limits_{n\to+\infty}} P^n(\sup\limits_{t\leq T} \mid< M^{n,\theta} >_t - \int\limits_0^t c^\theta(s, X^{n,\theta})du_s \mid\geq \epsilon) = 0.$

Conditions of group 2. For every $\theta, \theta' \in I\!R^m$, $\epsilon > 0$ and $T > 0$ we have:

5) $\overline{\lim\limits_{n\to+\infty}} P^n(\sup\limits_{t\leq T} \mid< X^{n,\theta,c}, X^{n,\theta',c} >_t - \int\limits_0^t d^{\theta,\theta'}(s, X^{n,\theta}, X^{n,\theta'})du_s \mid\geq \epsilon) = 0.$

Conditions of group 3. There exists $p \geq 2$ and $\alpha > m$ such that for all bounded stopping times τ and $\theta, \theta' \in K_i$, $\theta \neq \theta'$, $i \geq 1$, we have

6) $\sup\limits_n(E^n \mid A_\tau^{n,\theta} - A_\tau^{n,\theta'} \mid^p / \mid \theta - \theta' \mid^\alpha) \leq C_i,$

7) $\sup\limits_n(E(< X^{n,\theta,c} - X^{n,\theta',c} >_\tau)^{p/2} / \mid \theta - \theta' \mid^\alpha) \leq C_i,$

8) $\sup\limits_n(E\int\limits_0^\tau \int\limits_{\mathrm{R}^2\setminus\{(0,0)\}} (x-y)^2 d\nu^{n,\theta,\theta'} / \mid \theta - \theta' \mid^\alpha) \leq C_i,$

9) $\sup\limits_n(E\int\limits_0^\tau \int\limits_{\mathrm{R}^2\setminus\{(0,0)\}} \mid x-y \mid^p d\nu^{n,\theta,\theta'} / \mid \theta - \theta' \mid^\alpha) \leq C_i.$

Theorem 3 *We suppose that the conditions 1)-9) are satisfied. Then there are modifications of the processes $X^n = (X_t^{n,\theta})_{\theta\in\mathrm{R}^m, t\geq 0}$ and $X = (X_t^\theta)_{\theta\in\mathrm{R}^m, t\geq 0}$ with paths in $D(I\!R^+, C_{loc})$ such that a weak convergence*

$$X^n \stackrel{w(P^n)}{\longrightarrow} X$$

in $D(I\!R^+, C_{loc})$ takes place.

Proof. To prove our theorem we verify the conditions of the theorem 2 according to the following plan. We suppose that X^n and X are the paths in $D(I\!R^+, C_{loc})$. Using the conditions of group 1, we establish a weak convergence of the processes $X^{n,\theta}$ and X^θ in the Skorohod space $D(I\!R^+, I\!R)$. In turn, this convergence implies the condition 1) of theorem 2.

Then, using the conditions of groups 1 and 2 we establish the convergence of the finite-dimensionnal distributions

$$(X^{n,\theta_1}, X^{n,\theta_2}, \ldots, X^{n,\theta_l}) \stackrel{w(P^n)}{\longrightarrow} (X^{\theta_1}, X^{\theta_2}, \ldots, X^{\theta_l})$$

in the Skorohod space $D(I\!R^+, I\!R^l)$ for every $l \geq 2$. Finally, using the conditions of group 3 and the inequality (10) given below, we prove that for all $\theta, \theta' \in K_i$ and all bounded stopping times τ:

$$\sup_n E^n \mid X_\tau^{n,\theta} - X_\tau^{n,\theta'} \mid^p / \mid \theta - \theta' \mid^\alpha \leq \overline{C_i}.$$

Using specially choosen stopping time, this inequality and the lemma about the estimation of the modulus of continuity permit us to verify the condition 2) of theorem 2.

To prove a weak convergence of $X^{n,\theta}$ to X^θ for each $\theta \in I\!R^m$, in the Skorohod space $D(I\!R^+, I\!R)$, it is sufficient to verify the conditions of theorem 1 of Liptser, Shiryayev [21] p. 608-609 . In fact, since the "limit" semimartingale is continuous we have $\nu^\theta = 0$ and the conditions U_1 and U_2 of this theorem follow from the condition 2) of group 1.

The process $B^{n,\theta,a}$ appearing in the canonical decomposition of the semimartingale $X^{n,\theta}$ is such that

$$B_t^{n,\theta,a} = A_t^{n,\theta} + \int_0^t \int_{|x|>a} x d\nu^{n,\theta}$$

and condition 2) gives

$$\sup_{t \leq L} \mid B_t^{n,\theta,a} - A_t^{n,\theta} \mid \stackrel{P^n}{\longrightarrow} 0,$$

then we get condition U_3 after replacing $A_t^{n,\theta}$ by $B_t^{n,\theta,a}$.

The expression for the predictable variation of the square- integrable martingale $M^{n,\theta,a}$ appearing in the canonical decomposition of the semimartingale $X^{n,\theta}$ is :

$$< M^{n,\theta,a} >_t = < X^{n,\theta,c} >_t + \int_0^t \int_{|x|\leq a} x^2 d\nu^{n,\theta} - \sum_{0<s\leq t} (\int_{|x|\leq a} x \nu^{n,\theta}(\{s\}, dx))^2,$$

$$< M^{n,\theta} >_t = < X^{n,\theta,c} >_t + \int_0^t \int_{\mathrm{R}\setminus\{0\}} x^2 d\nu^{n,\theta} - \sum_{0<s\le t} (\int_{\mathrm{R}\setminus\{0\}} x\nu^{n,\theta}(\{s\},dx))^2,$$

So, the difference between these two predictable variations can be estimated as follows:

$$\begin{aligned} &|< M^{n,\theta} >_t - < M^{n,\theta,a} >_t| \\ &\le | \int_0^t \int_{|x|>a} x^2 d\nu^{n,\theta} - \sum_{0<s\le t} \int_{|x|>a} x\nu^{n,\theta}(\{s\},dx) \\ &\times [2\int_{|x|\le a} x\nu^{n,\theta}(\{s\},dx) + \int_{|x|>a} x\nu^{n,\theta}(\{s\},dx)]| \\ &\le 4\int_0^t \int_{|x|>a} x^2 d\nu^{n,\theta}, \end{aligned}$$

and by condition 2) the right-hand side of this inequality tends to zero in P^n-probability . This means that condition 3) is satisfied when we replace $< M^{n,\theta} >_t$ by $< M^{n,\theta,a} >_t$, i.e. the condition U_4 holds.

Finally, all the conditions of theorem 1 of Liptser, Shiryayev [21], p.608-609 are satisfied and we have

$$X^{n,\theta} \stackrel{w(P^n)}{\longrightarrow} X^\theta$$

for every $\theta \in I\!R^m$ in the Skorohod space $D(I\!R^+, I\!R)$.

To prove the convergence of the finite-dimensional distributions we show that

$$(X^{n,\theta_1}, X^{n,\theta_2}, \dots, X^{n,\theta_l}) \stackrel{w(P^n)}{\longrightarrow} (X^{\theta_1}, X^{\theta_2}, \dots, X^{\theta_l})$$

in the Skorohod space $D(I\!R^+, I\!R^l)$. In turn, according to the Kramer-Wold principle (see [2], p.), this convergence is equivalent to

$$\sum_{i=1}^l c_i X^{n,\theta_i} \stackrel{w(P^n)}{\longrightarrow} \sum_{i=1}^l c_i X^{\theta_i} \tag{8}$$

for every $c_i \in I\!R$ and $\theta_i \in I\!R^m$. Without loss of generality we can and will suppose that $c_i \ne 0$ for all $1 \le i \le l$.

We notice that the process $\sum_{i=1}^l c_i X^{n,\theta_i}$ with arbitrary $c_i \in I\!R$ and $\theta_i \in I\!R^m$ is a special locally square-integrable semimartingale with the decomposition

$$\sum_{i=1}^l c_i X_t^{n,\theta_i} = \sum_{i=1}^l c_i X_0^{n,\theta_i} + \sum_{i=1}^l c_i A_t^{n,\theta_i} + \sum_{i=1}^l c_i M_t^{n,\theta_i},$$

and $\sum_{i=1}^{l} c_i X^{\theta_i}$ is continuous with the following decomposition

$$\sum_{i=1}^{l} c_i X_t^{\theta_i} = \sum_{i=1}^{l} c_i X_0^{\theta_i} + \sum_{i=1}^{l} c_i A_t^{\theta_i} + \sum_{i=1}^{l} c_i M_t^{\theta_i}.$$

To prove a weak convergence (8) we can verify the conditions of group 1 for $\sum_{i=1}^{l} c_i X^{n,\theta_i}$ and $\sum_{i=1}^{l} c_i X^{\theta_i}$ and show that the triplet of the semimartingale $\sum_{i=1}^{l} c_i X^{\theta_i}$ satisfies the conditions of the type (2), (3), (4) and (5).

Condition 1) is clearly verified since

$$\sum_{i=1}^{l} c_i X_0^{n,\theta_i} \xrightarrow{P^n} \sum_{i=1}^{l} c_i X_0^{\theta_i}.$$

To verify condition 2) of group 1 we denote by $\nu^{n,\theta_1,\theta_2,\ldots,\theta_l}$ the compensator of the jump measure of $(X^{n,\theta_1}, \ldots, X^{n,\theta_l})$. Since

$$| \; c_1 x_1 + c_2 x_2 + \ldots + c_l x_l \; | \leq |x| \; |c|$$

where $x = ^{\top}(x_1, x_2, \ldots, x_l)$, $c = ^{\top}(c_1, c_2, \ldots, c_l)$, and $|.|$ is a euclidean norm, we have

$$\int_0^t \int_{|c_1x_1+c_2x_2+\ldots+c_lx_l|>a} | \; c_1 x_1 + c_2 x_2 + \ldots + c_l x_l \; |^2 \; d\nu^{n,\theta_1,\theta_2,\ldots,\theta_l}$$

$$\leq |c|^2 \int_0^t \int_{|x|>a/|c|} |x|^2 d\nu^{n,\theta_1,\theta_2,\ldots,\theta_l} \leq |c|^2 \sum_{i=1}^{l} \int_0^t \int_{|x|>a/|c|} x_i^2 d\nu^{n,\theta_1,\theta_2,\ldots,\theta_l} \qquad (9)$$

Noting that for every $a > 0$

$$\{x = ^{\top}(x_1, x_2, \ldots, x_l) : |x| > a\} \subset \bigcup_{i=1}^{l} \{x = ^{\top}(x_1, x_2, \ldots, x_l) : | \; x_i \; | > \frac{a}{l}\}$$

and we obtain

$$\int_0^t \int_{|x|>a} x_i^2 d\nu^{n,\theta_1,\theta_2,\ldots,\theta_l} \leq \sum_{j=1}^{l} \int_0^t \int_{|x_j|>a/l} x_i^2 d\nu^{n,\theta_1,\theta_2,\ldots,\theta_l}$$

$$\leq 2l \sum_{j=1}^{l} \int_0^t \int_{|x_j|>a/l} x_j^2 d\nu^{n,\theta_j}.$$

Since each term on the right-hand side of (9) tends to zero in P^n-probability, condition 2) of group 1 for the semimartingale $\sum_{i=1}^l c_i X^{n,\theta_i}$ is verified.

To obtain condition 3) of group 1 for $\sum_{i=1}^l c_i X^{n,\theta_i}$ we notice that

$$P^n(\sup_{t\leq L} \mid \sum_{i=1}^l c_i A_t^{n,\theta_i} - \sum_{i=1}^l c_i \int_0^t a^{\theta_i}(s, X^{n,\theta_i}) du_s \mid \geq \epsilon)$$
$$\leq \sum_{i=1}^l P^n(\sup_{t\leq L} \mid A_t^{n,\theta_i} - \int_0^t a^{\theta_i}(s, X^{n,\theta_i}) du_s \mid \geq \frac{\epsilon}{\mid c_i \mid l}),$$

and we use condition 3) of group 1 for each semimartingale X^{n,θ_i}. To verify the condition 4) for $\sum_{i=1}^l c_i X^{n,\theta_i}$ we write

$$< \sum_{i=1}^l c_i M^{n,\theta_i} >_t = \sum_{i=1}^l \sum_{i=1}^l c_i c_j < M^{n,\theta_i}, M^{n,\theta_j} >_t,$$

$$< \sum_{i=1}^l c_i M^{\theta_i} >_t = \sum_{i=1}^l \sum_{i=1}^l c_i c_j < M^{\theta_i}, M^{\theta_j} >_t,$$

so that condition 4) for $\sum_{i=1}^l c_i X^{n,\theta_i}$ follows from the corresponding conditions 4) and 5) of groups 1 and 2 respectively.

Now we verify the conditions of the type (2), (3), (4), (5), and (6) for the semimartingales $\sum_{i=1}^l c_i X^{\theta_i}$. These are satisfied because

$$\sum_{i=1}^l c_i A_t^{\theta_i}(X^i) = \int_0^t (\sum_{i=1}^l c_i a^{\theta_i}(s, X^i)) du_s,$$

$$< \sum_{i=1}^l c_i M^{\theta_i} >_t = \int_0^t (\sum_{i=1}^l c_i^2 c^{\theta_i}(s, X^i)) + \sum_{i\neq j} c_i c_j d^{\theta_i,\theta_j}(s, X^i, X^j)) du_s.$$

In addition we have

$$\mid \sum_{i=1}^l c_i A_t^{\theta_i}(X^i) \mid \leq \sum_{i=1}^l |c_i| L(t)(1 + \sup_{s<t} \mid X_s^i \mid) \leq (\sum_{i=1}^l |c_i| L(t))(1 + \sup_{s<t} |X_s|),$$

$$\mid \sum_{i=1}^l c_i^2 c^{\theta_i}(s, X^i)) + \sum_{i\neq j} c_i c_j d^{\theta_i,\theta_j}(s, X^i, X^j) \mid \leq (\sum_{i=1}^l |c_i|)^2 L(t)(1 + \sup_{s<t} |X_s|^2),$$

where $|X_s|$ is the euclidean norme in $I\!R^l$ of $X_s = ^\top(X_s^1, X_s^2, \cdots X_s^l)$.

For the third part of the proof we will derive the following estimation: there is a constant $c(p) > 0$ such that for every bounded $I\!F^n$-stopping time τ and $t \geq 0$

$$\begin{aligned} E^n \mid X^{n,\theta}_{\tau\wedge t} - X^{n,\theta'}_{\tau\wedge t} \mid^p \leq c(p)\{E^n \mid X^{n,\theta}_0 - X^{n,\theta'}_0 \mid^p \\ +E^n \mid A^{n,\theta}_{\tau\wedge t} - A^{n,\theta'}_{\tau\wedge t} \mid^p + E^n(< X^{n,\theta,c} - X^{n,\theta',c} >^{p/2}_{\tau\wedge t}) \\ +E^n[\int_0^{\tau\wedge t}\int_{\mathrm{R}^2\setminus\{0,0\}}(x-y)^2 d\nu^{n,\theta,\theta'}]^{p/2} + E^n\int_0^{\tau\wedge t}\int_{\mathrm{R}^2\setminus\{0,0\}}|x-y|^p d\nu^{n,\theta,\theta'}\}. \end{aligned} \quad (10)$$

From the semimartingale decomposition

$$X^{n,\theta}_{\tau\wedge t} = X_0 + A^{n,\theta}_{\tau\wedge t} + M^{n,\theta}_{\tau\wedge t}$$

and the inequality $(a+b)^p \leq 2^p(a^p + b^p)$, verified for all $a, b \geq 0$ and $p > 0$, we conclude that

$$\begin{aligned} E^n \mid X^{n,\theta}_{\tau\wedge t} - X^{n,\theta'}_{\tau\wedge t} \mid^p \leq c(p)\{E^n \mid X^{n,\theta}_0 - X^{n,\theta'}_0 \mid^p \\ +E^n \mid A^{n,\theta}_{\tau\wedge t} - A^{n,\theta'}_{\tau\wedge t} \mid^p + E^n \mid M^{n,\theta}_{\tau\wedge t} - M^{n,\theta'}_{\tau\wedge t} \mid^p\}. \end{aligned} \quad (11)$$

Since $(M^{n,\theta}_{\tau\wedge t} - M^{n,\theta'}_{\tau\wedge t})_{t\geq 0}$ is a locally square-integrable martingale, then the estimation of Valkeila, Dzaparidze [4] implies that for $p \geq 2$

$$\begin{aligned} E^n \mid M^{n,\theta}_{\tau\wedge t} - M^{n,\theta'}_{\tau\wedge t} \mid^p \leq c(p)\{E^n < M^{n,\theta} - M^{n,\theta'} >^{p/2}_{\tau\wedge t} \\ +E^n\int_0^{\tau\wedge t}\int_{\mathrm{R}^2\setminus\{0,0\}}|x-y|^p d\nu^{n,\theta,\theta'}\}, \end{aligned} \quad (12)$$

where c(p) is a constant, depending on p.

Noting that

$$\begin{aligned} &< M^{n,\theta} - M^{n,\theta'} >_{\tau\wedge t} = < X^{n,\theta,c} - X^{n,\theta',c} >_{\tau\wedge t} \\ &\quad + \int_0^{\tau\wedge t}\int_{\mathrm{R}^2\setminus\{0,0\}}(x-y)^2 d\nu^{n,\theta,\theta'} - \sum_{0<s\leq\tau\wedge t}(\int_{\mathrm{R}^2\setminus\{0,0\}}(x-y)\nu^{n,\theta,\theta'}(\{s\},dx,dy))^2 \\ &\quad \leq < X^{n,\theta,c} - X^{n,\theta',c} >_{\tau\wedge t} + 2\int_0^{\tau\wedge t}\int_{\mathrm{R}^2\setminus\{0,0\}}(x-y)^2 d\nu^{n,\theta,\theta'}. \end{aligned} \quad (13)$$

The relations (11), (12) and (13) imply the inequality (10) for some constant c(p) depending on p.

To verify condition 2) of theorem 2 we have to prove that

$$\lim_{h\to 0}\sup_n P^n(\sup_{0\le t\le T}\sup_{\substack{|\theta-\theta'|\le h\\ \theta,\theta'\in K_i}} \mid X_t^{n,\theta}-X_t^{n,\theta'}\mid\ge\epsilon)=0 \tag{14}$$

for all $\epsilon>0$ and $T>0$. For this we introduce a stopping time

$$\tau=\inf\{0\le t\le T:\sup_{\substack{|\theta-\theta'|\le h\\ \theta,\theta'\in K_i}} \mid X_t^{n,\theta}-X_t^{n,\theta'}\mid\ge\epsilon\}$$

with the convention that $inf(\emptyset)=T$. Then

$$P^n(\sup_{0\le t\le T}\sup_{\substack{|\theta-\theta'|\le h\\ \theta,\theta'\in K_i}} \mid X_t^{n,\theta}-X_t^{n,\theta'}\mid\ge\epsilon)\le P^n(\sup_{\substack{|\theta-\theta'|\le h\\ \theta,\theta'\in K_i}} \mid X_\tau^{n,\theta}-X_\tau^{n,\theta'}\mid\ge\epsilon). \tag{15}$$

To estimate the right-hand side of the inequality (15) we use the lemma 7.4 p. 1000 of Pfaff [19]. We then note that the conditions of group 3 and the inequality (10) give

$$E^n\mid X_{\tau\wedge t}^{n,\theta}-X_{\tau\wedge t}^{n,\theta'}\mid^p\le c(p)\mid\theta-\theta'\mid^\alpha$$

for every bounded $I\!\!F^n$-stopping time τ and $t\ge 0$. From this lemma it follows that

$$P^n(\sup_{\substack{|\theta-\theta'|\le h\\ \theta,\theta'\in K_i}} \mid X_{\tau\wedge t}^{n,\theta}-X_{\tau\wedge t}^{n,\theta'}\mid\ge\epsilon)\le c(p,\epsilon)h^{\alpha-m} \tag{16}$$

where $c(p,\epsilon)$ is some constant independent on n. Taking $\sup_n$ and $\lim_{h\to 0}$ and choosing t big enough to ensure that $\tau\wedge t=\tau$, we obtain condition 2) of theorem 2.

It remains to show how one can construct the modifications of the processes X^n and X with values in $D(I\!\!R^+,C_{loc})$. Repeating the same arguments as above, but with the rationals θ and θ', we obtain the estimation (16) and, hence, (14). Then taking the rationals $\theta\in Q(I\!\!R^m)$ we set for $\theta_0\in I\!\!R^m\setminus Q(I\!\!R^m)$ and $t\ge 0$

$$X_t^{n,\theta_0}=\lim_{\theta\to\theta_0}X_t^{n,\theta}.$$

Finally, we verify that for each $\epsilon > 0$, $\theta_0 \in I\!R^m$ and $N \in I\!N$

$$\lim_{h \to 0} P^n(W_h^N(X^{n,\theta_0}) \geq \epsilon) = 0. \tag{17}$$

In fact, this is true for $\theta_0 \in Q(I\!R^m)$, and for irrational θ_0 we choose $\theta \in Q(I\!R^m)$ such that $\mid \theta_0 - \theta \mid \leq h$ and write

$$W_h^N(X^{n,\theta_0}) \leq W_h^N(X^{n,\theta}) + 2V_h^{N,i}(X^n).$$

This inequality together with (14) and (17) gives the result. □

4 Consistency and weak convergence of the minimal distance estimators.

We suppose that on the filtered probability space $(\Omega, \mathcal{F}, I\!F, P^\epsilon)$ we are given a sequence of special square-integrable semimartingales $X^{\epsilon,\theta} = (X_t^{\epsilon,\theta})_{t\geq 0}$ depending on the parameter $\theta \in \Theta \subset I\!R^m$ and having the decomposition (1). We suppose that Θ is bounded closed convex set.

We assume that $X^\theta = (X_t^\theta)_{t\geq 0}$ is a deterministic function with paths in $D(I\!R^+, I\!R)$ having a locally bounded variation. This function will play the role of the "limit" trajectory as $\epsilon \to 0$ of the process $X^{\epsilon,\theta}$.

We suppose that the space of paths $D([0,T], I\!R)$ with $T > 0$ is endowed with a norm $||.||_T$,bounded above by the corresponding supremum norm and that for every function $X \in D(I\!R^+, I\!R)$, $||X||_T$ is a non-decreasing right-continuous function of T.

We suppose further that the "limit" paths are distinguishable in the norm $||\,.\,||$ i.e. for all $T > 0, \theta, \theta' \in \Theta, \theta \neq \theta'$ we have

$$||X_.^\theta - X_.^{\theta'}||_T > 0,$$

and that they are continuous in this norm, i.e. $||X_.^\theta - X_.^{\theta'}||_T$ tends to zero as $\theta' \to \theta$ for every $\theta \in \Theta$.

We assume that the true value of the parameter is θ_0 , $\theta_0 \in \mathrm{int}(\Theta)$, and we define the minimal distance estimator $\hat{\theta}_T^\epsilon$ corresponding to observations on the interval $[0,T]$ as

$$\hat{\theta}_T^\epsilon = \arg \inf_{\theta \in \Theta} ||X^{\epsilon,\theta_0} - X^\theta||_T.$$

In the following theorem we give conditions under which the estimator $\hat{\theta}_T^\epsilon$ is consistent .

Theorem 4 *We assume that for $T > 0$ as $\epsilon \to 0$*

1) $||A^{\epsilon,\theta_0} - X^{\theta_0}||_T \xrightarrow{P^\epsilon} 0,$

2) $< M^{\epsilon,\theta_0} >_T \xrightarrow{P^\epsilon} 0.$

Then the estimator $\hat{\theta}_T^\epsilon$ is a consistent estimator of θ_0.

Proof. We consider the random elements: $Y_T^{\epsilon,\theta_0} = ||X^{\epsilon,\theta_0} - X^\theta||_T$ and $Y_T^{\theta_0} = ||X^{\theta_0} - X^\theta||_T$. It is easy to see that for each θ, θ' we have

$$|Y_T^{\epsilon,\theta} - Y_T^{\epsilon,\theta'}| \leq ||X^\theta - X^{\theta'}||_T, \quad |Y_T^\theta - Y_T^{\theta'}| \leq ||X^\theta - X^{\theta'}||_T$$

and, since in the norm $||.||_T$ the limit paths are continuous in θ , $Y_T^{\epsilon,\theta}$ and Y_T^θ are in $C(\Theta)$.

We will now show that the laws of $Y_T^\epsilon = (Y_T^{\epsilon,\theta})_{\theta\in\Theta}$ converge in $C(\Theta)$ to the law of $Y_T = (Y_T^\theta)_{\theta\in\Theta}$. For this, using a semimartingale decomposition, we can write that

$$\sup_{\theta\in\Theta} |Y_T^{\epsilon,\theta} - Y_T^\theta| \leq ||X^{\epsilon,\theta_0} - X^{\theta_0}||_T \leq \tag{18}$$
$$\leq ||A^{\epsilon,\theta_0} - X^{\theta_0}||_T + ||M^{\epsilon,\theta_0}||_T \leq ||A^{\epsilon,\theta_0} - X^{\theta_0}||_T + \sup_{0\leq t\leq T} |M_t^{\epsilon,\theta_0}|$$

From the Lenglart-Rebolledo inequality ([21], p. 46)we have for each $a > 0, b > 0$

$$P^\epsilon(\sup_{0\leq t\leq T} |M_t^{\epsilon,\theta_0}| > a) \leq b/a^2 + P^\epsilon(< M^{\epsilon,\theta_0} >_T > b). \tag{19}$$

This inequality (19) together with (18) and the conditions 1) and 2) of the theorem imply that

$$\sup_{\theta\in\Theta} |Y_T^{\epsilon,\theta} - Y_T^\theta| \xrightarrow{P^\epsilon} 0$$

which establishes the needed convergence.

Now consider the functional ϕ such that

$$\phi(Y_T^\epsilon) = arg \inf_{\theta\in\Theta} Y_T^{\epsilon,\theta}.$$

From the condition of distinguishability we see that $\phi(Y_T) = arg \inf_{\theta\in\Theta} ||X^{\theta_0} - X^{\theta}||_T = \theta_0$ is unique . Hence, ϕ is continuous and the estimator

$$\hat{\theta}_T^{\epsilon} = arg \inf_{\theta\in\Theta} ||X^{\epsilon,\theta_0} - X^{\theta}||_T$$

converges to θ_0 in law, and, also in probability. □

To study a weak convergence of normalized minimum distance estimators

$$\hat{\beta}_T^{\epsilon} = \frac{1}{\epsilon}(\hat{\theta}_T^{\epsilon} - \theta_0)$$

we introduce a new parameter $\beta = (\theta - \theta_0)/\epsilon$ and set $\tilde{X}^{\epsilon,\theta_0} = (\tilde{X}_t^{\epsilon,\theta_0})_{t\geq 0}$ with

$$\tilde{X}_t^{\epsilon,\theta_0} = (X_t^{\epsilon,\theta_0} - X_t^{\theta_0})/\epsilon.$$

We also introduce a continuous gaussian process $U^{\theta_0} = (U_t^{\theta_0})_{t\geq 0}$ playing the role of the limit process for $\tilde{X}^{\epsilon,\theta_0}$, and having the decomposition

$$U_t^{\theta_0} = B_t^{\theta_0} + N_t^{\theta_0}$$

where B^{θ_0} is the predictable process of locally integrable variation and N^{θ_0} is a local martingale.

We suppose that the solution of the corresponding semimartingale problem is unique and that there exist predictable functions $b^{\theta_0}(t, X)$, $c^{\theta_0}(t, X)$ with the modulus integrable with respect to some continuous increasing deterministic function (u_s) of locally bounded variation such that

$$B_t^{\theta_0}(X) = \int_o^t b^{\theta_0}(s, X) du_s,$$

$$< N^{\theta_0} >_t (X) = \int_o^t c^{\theta_0}(s, X) du_s.$$

We assume that the functions $b^{\theta_0}(t, X)$ and $c^{\theta_0}(t, X)$ are continuous in the Skorohod metric and they verify

$$| b^{\theta_0}(t, X) | \leq L(t)(1 + \sup_{s<t} |X_s|),$$

$$| c^{\theta_0}(t, X) | \leq L(t)(1 + \sup_{s<t} |X_s|^2),$$

where L(t) is positive and integrable on the finite intervals function which can depend on θ_0 .

Theorem 5 *We suppose that for each $\delta > 0$ and $T > 0$*

1) $(X_0^{\epsilon,\theta_0} - X_0^{\theta_0})/\epsilon \xrightarrow{P^\epsilon} 0,$

2) $\overline{\lim_{\epsilon\to 0}}\ P^\epsilon(\int_0^T \int_{|x|>\gamma} |x|^2 d\tilde{\nu}^{\epsilon,\theta_0} \geq \delta) = 0, \ \ \forall \gamma > 0,$

3) $\overline{\lim_{\epsilon\to 0}}\ P^\epsilon(\sup_{t\leq T} |\frac{A_t^{\epsilon,\theta_0} - X_t^{\theta_0}}{\epsilon} - \int_0^t b(s, \tilde{X}^{\epsilon,\theta_0}) du_s| \geq \delta) = 0,$

4) $\overline{\lim_{\epsilon\to 0}}\ P^\epsilon(\sup_{t\leq T} |\frac{< M^{\epsilon,\theta_0} >_t}{\epsilon^2} - \int_0^t c(s, \tilde{X}^{\epsilon,\theta_0}) du_s| \geq \delta) = 0,$

5) there exists a function $\overset{\circ}{X}{}^{\theta_0} \in D(I\!R^+, I\!R)$ *such that as* $|\Delta| \to 0$

$$\frac{||X^{\theta_0+\Delta} - X^{\theta_0} - {}^\top\Delta \overset{\circ}{X}{}^{\theta_0} ||_T}{|\Delta|} \to 0,$$

6) the strong identifiability condition that

$$\underline{\lim}_{L\to+\infty} \ \ \underline{\lim}_{\epsilon\to 0} \ \ \inf_{|\beta|>L} ||\frac{1}{\epsilon}(X^{\theta_0+\epsilon\beta} - X^{\theta_0})||_T = +\infty,$$

7) $arg \inf_{\beta\in \boldsymbol{R}^m} ||U^{\theta_0} - {}^\top\beta \overset{\circ}{X}{}^{\theta_0} ||_T$ *is unique .*

Then there is a weak convergence of the normalized estimators $\hat{\beta}_T^\epsilon$ *to the estimator* $\hat{\beta}_T$ *defined by*

$$\hat{\beta}_T = arg \inf_{\beta\in \boldsymbol{R}^m} ||U^{\theta_0} - {}^\top\beta \overset{\circ}{X}{}^{\theta_0} ||_T$$

Remark 1. To clarify the sense of the conditions of this theorem, we introduce the processes $Y^{\epsilon,\beta} = (Y_t^{\epsilon,\beta})_{t\geq 0}$ with

$$Y_t^{\epsilon,\beta} = ||\frac{1}{\epsilon}(X^{\epsilon,\theta_0} - X^{\theta_0+\epsilon\beta})||_t \tag{20}$$

and the process $Y^{\beta} = (Y_t^{\beta})_{t\geq 0}$ with

$$Y_t^{\beta} = ||U^{\theta_0} -^{\top}\beta \overset{\circ}{X}{}^{\theta_0} ||_t. \tag{21}$$

One can see that the conditions 1) - 5) assure a weak convergence of the processes $Y^{\epsilon} = (Y^{\epsilon,\beta})_{\beta\in R^m}$ to the process $Y = (Y^{\beta})_{\beta\in R^m}$ in the space $D(I\!R^+, C_{loc})$. Condition 1) assures the convergence of the initial values, condition 2) is a sort of Lindenberg condition and conditions 3) and 4) are the conditions of differentiability of A^{ϵ,θ_0} and $< M^{\epsilon,\theta_0} >$ in the sense of the probability distance furnished with the sup in t norm. In turn, conditions 6) and 7) assure the continuity of the functional $arg\inf$.

We continue with a discussion of the strong identifiability condition.

Lemma 1 *We suppose that the limit trajectory X^{θ} is differentiable in θ_0 with respect to the norm $||.||_T$. Then the following two conditions are equivalent:*

1) the strong identifiability condition that

$$\lim_{L\to+\infty} \lim_{\epsilon\to 0} \inf_{|\beta|>L} ||\frac{1}{\epsilon}(X^{\theta_0+\epsilon\beta} - X^{\theta_0})||_T = +\infty,$$

2) the identifiability condition that for every $\nu > 0$, $\inf_{|\theta-\theta_0|>\nu} ||X^{\theta} - X^{\theta_0}||_T > 0$ and the non-singularity condition that there exists $c > 0$ such that for all $b \in I\!R^m$, $||(\overset{\circ}{X}{}^{\theta_0}, b)||_T \geq c|b|$.

Proof. Suppose that condition 1) is satisfied and the non-singularity condition is not . Then we can find a sequence of b_n with $|b_n| = 1$ such that $||(\overset{\circ}{X}{}^{\theta_0}, b_n)||_T \to 0$ as $n \to \infty$. Let $L > 0$. Then there exists n_0 such that for all $n \geq n_0$ we have $||(\overset{\circ}{X}{}^{\theta_0}, b_n)||_T \leq 1/L$. Condition 1) applied to $\beta = 2Lb_n$ gives

$$\lim_{L\to+\infty} 2L \lim_{\epsilon\to 0} \frac{||X^{\theta_0+2Lb_n\epsilon} - X^{\theta_0}||_T}{2L\epsilon} = +\infty$$

but

$$L \lim_{\epsilon\to 0} \frac{||X^{\theta_0+2Lb_n\epsilon} - X^{\theta_0}||_T}{2\epsilon L} = ||(\overset{\circ}{X}{}^{\theta_0}, b_n)||_T \leq 1$$

for sufficiently big n , contradicting with condition 1).

If we suppose that there exists ν_0 such that $\inf_{|\theta-\theta_0|>\nu_0} ||X^\theta - X^{\theta_0}||_T = 0$, then for every $L > 0$ there exists ϵ_0 such that for all $0 < \epsilon < \epsilon_0$ we have $L\epsilon \leq \nu_0$ and

$$\inf_{|\beta|>L} ||X^\theta - X^{\theta_0}||_T \leq \inf_{|\theta-\theta_0|>\nu_0} ||X^\theta - X^{\theta_0}||_T = 0$$

which also contradicting condition 1).

We suppose now that the conditions 2) are satisfied and we fix $L > 0$. By the differentiability condition we can find $\nu_0 > 0$ such that for all $|\theta-\theta_0| \leq \nu_0$ we have

$$||X_\theta - X_{\theta_0} -^\top (\theta-\theta_0)\ \overset{\circ}{X}{}^{\theta_0}\ ||_T \leq 1/2c|\theta-\theta_0|,$$

where $c > 0$ is taken from the non-singularity condition. By the condition of identifiability, we have $\inf_{|\theta-\theta_0|>\nu} ||X^\theta - X^{\theta_0}||_T > 0$ and by the continuity of X_θ in θ_0, we can find $0 < \delta < \nu_0$ such that for all $|\theta-\theta_0| \leq \delta$ we have

$$||X^\theta - X^{\theta_0}||_T < \inf_{|\theta-\theta_0|>\nu_0} ||X^\theta - X^{\theta_0}||_T.$$

Choose ϵ_0 such that for all $0 < \epsilon < \epsilon_0$ we have $\epsilon L < \delta$,then for such ϵ we obtain that

$$\inf_{|\beta|>L} ||\frac{1}{\epsilon}(X^{\theta_0+\epsilon\beta} - X^{\theta_0})||_T = \frac{1}{\epsilon} \inf_{|\theta-\theta_0|>L\epsilon} ||X^\theta - X^{\theta_0}||_T$$

$$= \frac{1}{\epsilon} \inf_{\epsilon L<|\theta-\theta_0|\leq\nu_0} ||X^\theta - X^{\theta_0}||_T = \inf_{L<|\beta|\leq\nu_0/\epsilon} ||\frac{1}{\epsilon}(X^{\theta_0+\beta\epsilon} - X^{\theta_0})||_T$$

$$\geq \frac{1}{\epsilon} \inf_{L<|\beta|\leq\nu_0/\epsilon} (||(\overset{\circ}{X}{}^{\theta_0}, \beta\epsilon)||_T - ||X_\theta - X_{\theta_0} -^\top (\theta-\theta_0)\ \overset{\circ}{X}{}^{\theta_0}\ ||_T) \geq 1/2cL$$

which tends to infinity as $L \to \infty$. □

To prove our theorem we start with a discussion of the relation between the convergence in $D(I\!R^+, C_{loc})$ and the continuity of the functional $arg \inf$.

We denote by $D(I\!R^+, C^0_{loc})$ the subspace of the space $D(I\!R^+, C_{loc})$ containing the elements Z such that for all continuity points $1/N$ and N we have

$$\sup_{1/N \le t \le N} \sup_{|\beta|>L} |Z_t(\beta)| \to 0$$

as $L \to \infty$. For $Z \in D(I\!R^+, C^0_{loc})$ and a point of continuity $T > 0$ of Z we set

$$\hat{\beta}^n_T = arg \sup_{\beta \in \boldsymbol{R}^m} |Z^n_T(\beta)| \quad \hat{\beta}_T = arg \sup_{\beta \in \boldsymbol{R}^m} |Z_T(\beta)|.$$

We suppose as usual that if $\sup_{\beta \in \boldsymbol{R}^m} |Z^n_T(\beta)|$ is achieved in several points we choose one of them arbitrarily. We note that $\sup_{\beta \in \boldsymbol{R}^m} |Z^n_T(\beta)|$ can be not achieved. In this case we put $arg \sup_{\beta \in \boldsymbol{R}^m} |Z^n_T(\beta)| = +\infty$ and $|Z^n_T(+\infty)| = \overline{\lim}_{|\beta| \to +\infty} |Z^n_T(\beta)|$.

Lemma 2 *We suppose that the sequence of elements* $Z^n = (Z^n_t(\beta))_{\beta \in \boldsymbol{R}^m, t \ge 0}$ *from* $D(I\!R^+, C_{loc})$ *converges in the Skorohod topology to an element* $Z = (Z_t(\beta))_{\beta \in \boldsymbol{R}^m, t \ge 0}$ *from* $D(I\!R^+, C^0_{loc})$. *We suppose also that for a point of continuity* T *of* Z *we have:*

$$\lim_{L \to +\infty} \overline{\lim_{n \to +\infty}} \sup_{|\beta|>L} |Z^n_T(\beta)| = 0. \tag{22}$$

Then, if $\hat{\beta}_T$ *is unique, we have*

$$\hat{\beta}^n_T \to \hat{\beta}_T.$$

Remark 2 One easily sees that there are examples where the condition of lemma 1 is not satisfied and the convergence of $\hat{\beta}^n_T$ fails.

Proof. If we suppose that $\hat{\beta}^n_T$ does not converge to $\hat{\beta}_T$ then we can have two cases:

a) there exists a compact set $\mathcal{K}$ such that $\hat{\beta}^n_T \in \mathcal{K}$ for each $n \ge 1$;

b) there exists a subsequence (n) such that $|\hat{\beta}^n_T| \to +\infty$.

In case a) one easily gets a contradiction with the convergence of Z^n to Z in $D(I\!R^+, C_{loc})$ and the uniqueness of $\hat{\beta}_T$. In the case b) we write for each $L > 0$

$$\overline{\lim_{n \to \infty}} |Z^n_T(\hat{\beta}^n_T) - Z_T(\hat{\beta}^n_T)| \le \overline{\lim_{n \to +\infty}} \sup_{|\beta|>L} |Z^n_T(\beta)| + \sup_{|\beta|>L} |Z_T(\beta)|$$

and then letting $L \to +\infty$ and using condition (22) we get that

$$\overline{\lim_{n\to\infty}} |Z_T^n(\hat{\beta}_T^n) - Z_T(\hat{\beta}_T^n)| = 0.$$

Since $Z \in D(I\!R^+, C^0_{loc})$ this leads to

$$\overline{\lim_{n\to\infty}} Z_T^n(\hat{\beta}_T^n) = 0.$$

But $|Z_T^n(\hat{\beta}^n)| = \sup_{\beta\in \boldsymbol{R}^m} |Z_T^n(\beta)|$ and the convergence in $D(I\!R^+, C_{loc})$ gives $\sup_{\beta\in K_i} |Z_T^n(\beta)| \to \sup_{\beta\in K_i} |Z_T(\beta)|$ for each $i \geq 1$. This implies that $Z_T(\beta) = 0$ for all $\beta \in I\!R^m$, contradicting the unicity of $\hat{\beta}_T$. $\square$

Lemma 3 *We suppose that Z^n is a sequence of random processes with paths in $D(I\!R^+, C_{loc})$ converging weakly to $Z \in D(I\!R^+, C^0_{loc})$. Let T be a point of continuity of Z. If $\hat{\beta}_T$ is unique and for every $\delta > 0$ we have*

$$\lim_{L\to+\infty} \overline{\lim_{n\to+\infty}} P(\sup_{|\beta|>L} |Z_T^n(\beta)| \geq \delta) = 0, \qquad (23)$$

then $\hat{\beta}_T^n$ converges weakly to $\hat{\beta}_T$.

Proof. The application $\phi(Z) = arg \sup_{\beta\in \boldsymbol{R}^m} |Z(\beta)|$ with the usual convention concerning $arg \sup$ as given above, is continuous on the set

$$A = \{Z \in C_{loc} : \lim_{L\to+\infty} \sup_{|\beta|>L} |Z(\beta)| = 0\}.$$

Denoting by Q^n and Q the laws on C_{loc} of Z_T^n and Z_T respectively. If $Q(A^c) = 0$ we have $\phi(Z_T^n) \to \phi(Z_T)$ as $n \to \infty$ weakly, but $Q(A^c) = 0$ is equivalent to

$$Q(\lim_{L\to\infty} \sup_{|\beta|>L} |Z(\beta)| > \delta) = 0,$$

for every $\delta > 0$, so we have

$$Q(\lim_{L\to\infty} \sup_{|\beta|>L} |Z(\beta)| > \delta) = \lim_{L\to\infty} Q(\sup_{|\beta|>L} |Z(\beta)| > \delta)$$

$$\leq \lim_{L\to\infty} \overline{\lim_{n\to\infty}} Q^n(\sup_{|\beta|>L} |Z(\beta)| > \delta) = 0.$$

□

Proof of theorem 5. To prove a weak convergence of the processes Y^ϵ to Y in $D(I\!R^+, C_{loc})$, as defined in Remark 1 by (20) and (21) respectively, we cannot directly apply the semimartingale convergence theorem, because for predictable continuous processes the martingale problem does not have a unique solution. This is why we start with the proof of a weak convergence of $U^\epsilon = (U_t^{\epsilon,\beta})_{\beta \in \boldsymbol{R}^m, t\geq 0}$ to $U = (U_t^\beta)_{\beta \in \boldsymbol{R}^m, t \geq 0}$ in $D(I\!R^+, C_{loc})$ where

$$U_t^{\epsilon,\beta} = \frac{1}{\epsilon}(X_t^{\epsilon,\theta_0} - X_t^{\theta_0}) - {}^\top\beta \overset{\circ}{X}{}_t^{\theta_0} \text{ and } U_t^\beta = U_t^{\theta_0} - {}^\top\beta \overset{\circ}{X}{}_t^{\theta_0} .$$

Since

$$U^{\epsilon,\beta} = \frac{1}{\epsilon}(A^{\epsilon,\theta_0} - X^{\theta_0}) - {}^\top\beta \overset{\circ}{X}{}^{\theta_0} + \frac{1}{\epsilon}M^{\epsilon,\theta_0}$$

conditions 1) - 4) imply the conditions of group 1 in theorem 3. The condition of group 2 is automatically satisfied since the martingale parts are independ on β. The conditions of group 3 are also satisfied since for all $0 \leq t \leq T$ we have

$$|\frac{1}{\epsilon}(A_t^{\epsilon,\theta_0} - X_t^{\theta_0}) - {}^\top\beta \overset{\circ}{X}{}_t^{\theta_0} - (\frac{1}{\epsilon}(A_t^{\epsilon,\theta_0} - X_t^{\theta_0}) - {}^\top\beta' \overset{\circ}{X}{}_t^{\theta_0})| \leq |\beta - \beta'| \sup_{0\leq t\leq T} | \overset{\circ}{X}{}_t^{\theta_0} |.$$

Since the norm $||.||$ is bounded above by the supremum norm, using the Skorohod representation theorem we see that the processes $||U^\epsilon|| = (||U^{\epsilon,\beta}||_t)_{\beta\in\boldsymbol{R}^m, t\geq 0}$ converges weakly to $||U|| = (||U^\beta||_t)_{\beta \in \boldsymbol{R}^m, t\geq 0}$ in the space $D(I\!R^+, C_{loc})$.

From condition 5) we have for all $0 < t \leq T$ and $\beta \in K_i$ that

$$| \, ||\frac{1}{\epsilon}(X^{\epsilon,\theta_0} - X^{\theta_0+\beta\epsilon})||_t - ||U^{\epsilon,\theta_0} - {}^\top\beta \overset{\circ}{X}{}^{\theta_0} ||_t \, | \leq ||\frac{1}{\epsilon}(X^{\theta_0+\beta\epsilon} - X^{\theta_0}) - {}^\top\beta \overset{\circ}{X}{}^{\theta_0} ||_T \to 0, \tag{24}$$

as $\epsilon \to 0$ and hence we have the weak convergence of Y^ϵ to Y in $D(I\!R^+, C_{loc})$ as defined in Remark 1 by (20) and (21) respectively.

To prove the convergence of $\hat\beta_T^\epsilon$ to $\hat\beta_T$ we use the Skorohod representation theorem and lemma 2 with the process Z^ϵ defined by : for $t \geq 0$

$$Z_t^{\epsilon,\beta} = \frac{1}{1 + Y_t^{\epsilon,\beta}}.$$

We now establish that (23) holds. We note that for all $|\beta| \geq L$ and for all $\epsilon \leq \epsilon_0(L)$

$$Y_T^{\epsilon,\beta} \geq \frac{1}{\epsilon}||X^{\theta_0+\epsilon\beta} - X^{\theta_0}||_T - \frac{1}{\epsilon}||X^{\epsilon,\theta_0} - X^{\theta_0}||_T$$

$$\geq \inf_{|\beta|\geq L} \frac{1}{\epsilon}||X^{\theta_0+\epsilon\beta} - X^{\theta_0}||_T - \frac{1}{\epsilon}||X^{\epsilon,\theta_0} - X^{\theta_0}||_T,$$

implying that

$$P^\epsilon(\sup_{|\beta|\geq L} Z_T^{\epsilon,\beta} \geq \delta) = P^\epsilon(\inf_{|\beta|\geq L} Y_T^{\epsilon,\beta} \leq \frac{1}{\delta} - 1)$$

$$\leq P^\epsilon(\frac{1}{\epsilon}||X^{\epsilon,\theta_0} - X^{\theta_0}||_T \geq \inf_{|\beta|\geq L} \frac{1}{\epsilon}||X^{\theta_0+\epsilon\beta} - X^{\theta_0}||_T - \frac{1}{\delta} + 1).$$

Choosing δ to have the points of continuity of the law of $||U^{\theta_0}||_T$ and letting $\epsilon \to 0, L \to +\infty$ we obtain (23) using the strong identifiability condition .

From a weak convergence of Y^ϵ to Y in $D(I\!R^+, C_{loc})$, the condition (23) and the uniqueness of $\hat{\beta}_T$ we get a weak convergence of $\hat{\beta}_T^\epsilon$ to $\hat{\beta}_T$. $\square$

Corollary 1 *(cf. Kutoyants([13]) Let the conditions 1)-6) of theorem 5 be satisfied and let* $||.||$ *be an* L_2*-norm. We suppose that the matrix* $V_T^{\theta_0} = \int_0^T \overset{\circ}{X}{}_s^{\theta_0} \; {}^\top\overset{\circ}{X}{}_s^{\theta_0} \, ds$ *is invertible. Then*

$$\hat{\beta}_T = (V_T^{\theta_0})^{-1} \int_0^T U_s^{\theta_0} \; \overset{\circ}{X}{}_s^{\theta_0} \, ds$$

and the law of $\hat{\beta}_T^\epsilon$ *is asymptotically gaussian* $\mathcal{N}(m_T, \sigma_T^2)$ *with*

$$m_T = (V_T^{\theta_0})^{-1} \int_0^T EB_s^{\theta_0} \; \overset{\circ}{X}{}_s^{\theta_0} \, ds,$$

$$\sigma_T^2 = (V_T^{\theta_0})^{-1}[\int_0^T \int_0^T E(N_s^{\theta_0} \; N_u^{\theta_0}) \; \overset{\circ}{X}{}_s^{\theta_0} \; {}^\top\overset{\circ}{X}{}_u^{\theta_0} \, dsdu \;](V_T^{\theta_0})^{-1}.$$

We now consider the Ornstein-Uhlenbeck process defined by the equation

$$dX_t = \theta X_t dt + \epsilon dW_t,$$

where $W = (W_t)_{t\geq 0}$ is a standard Wiener process, $t \in [0,T]$, $X_0 = x_0 \neq 0$ and $\theta \in \Theta$, Θ is a compact convex subset of $I\!R$. One can easily show that $X_t^\theta = x_0 \exp(\theta t)$ and that

$$A_t^{\epsilon,\theta} = \int\limits_0^t \theta X_s ds, \quad < M^{\epsilon,\theta} >_t = < \epsilon W >_t = \epsilon^2 t.$$

In this case the conditions of theorem 4 are satisfied and the conditions 1)-4) of theorem 5 are also satisfied with $b(t,X) = \theta_0 X_t$ and $c(t,X) = t$. Conditions 5), 6) are automatically satisfied since for $\theta \in \Theta$, $||x_0 t \exp(\theta t)||_T > 0$. Since

$$\hat{\beta}_T = arg \inf_{\beta \in \boldsymbol{R}^m} ||W_t - \beta x_0 t \exp(\theta_0 t)||_T,$$

condition 7) is satisfied for instance, for sup- and L^p-norms with $p \geq 1$.(see Henaff[6]).

Corollary 2 *(cf. Henaff[6], Kutoyants, Pilibossian [17]) In the case of the Ornstein-Uhlenbeck process the minimum distance estimators are consistent. If $\hat{\beta}_T$ is unique, then the normalised minimum distance estimators converge weakly to $\hat{\beta}_T$. If in addition we take the L^2-norm, the normalised minimum distance estimators will be asymptotically $\mathcal{N}(0, \sigma_T^2)$ where*

$$\sigma_T^2 = \left(\int\limits_0^T \int\limits_0^T st(t \wedge s) \exp((s+t)\theta_0) ds dt \right) \left(x_0 \int\limits_0^T s^2 \exp(2s\theta_0) ds \right)^{-2}.$$

References

[1] P. Bertrand and Yu.A. kutoyants. *A minimum distance estimator for partially observed linear stochastic systems.* Preprint 95-1, 1995.

[2] P. Billingsley. *Convergence of probability measures.* Wiley, New-York, 1968.

[3] E. Bolthausen. *Convergence in distribution of minimum distance estimators.* Metrika 24, p. 215-227, 1977.

[4] K. Dzhaparidze et E. Valkeila. *On the Hellinger type distances for filtered experiments.* Probability Theory and related Fields 85, p. 105-117, 1990.

[5] E. Fournie and Yu.A. Kutoyants. *Estimateur de la distance minimale pour processus de diffusion ergodiques.* INRIA, Rapports de recherche N 1952, p. 1-33, juillet 1993.

[6] S. Henaff *On minimum distance estimate of the parameter of the Ornstein-Uhlenbeck process.* Preprint 95-12, 1995.

[7] R. Höpfner and Yu.A. Kutoyants. *On minimum distance estimation in recurrent Markov step process I and II.* (submitted for publication).

[8] J. Jacod *Calcul stochastique et problème de martingales.* Lecture Notes in Mathematics 714, Springer-Verlag, 1979.

[9] J. Jacod et P. Mano. *Une évaluation de la distance entre les lois d'une semi-martingale et d'un processus à accroissements indépendants.* Stochastics, Vol. 25, p. 87-124, 1988.

[10] J. Jacod and A.N. Shiryayev. *Limit theorems for stochastic processes.* Springer-Verlag, Berlin, 1987.

[11] H. Koul and T. De Wet. *Minimum distance estimation in a linear regression model.* The Annals of Statistics, Vol. 11, $n°3$, p. 921-932, 1983.

[12] YU.A. KUTOYANTS. *Minimum distance parameter estimation for diffusion type observations.* C. R. Acad. Sci. Paris, 312, série I, p. 637-642, 1991.

[13] YU.A. KUTOYANTS. *Identification of dynamical systems with small noise.* Kluver Academic Publishers, London, 1995.

[14] YU.A. KUTOYANTS AND O. LESSI. *Minimum distance estimation for diffusion random fields.* Preprint, 1995.

[15] YU.A. KUTOYANTS AND F. LIESE. *On minimum distance estimation for spatial Poisson processes.* Ann. Academiae Scient. Fennicae, ser. A,I, Vol. 17, p. 65-71, 1992.

[16] YU.A. KUTOYANTS AND T. MOURID. *Estimation dans un modèle autoregressif avec retards.* C.R.Acad. Paris, 315, sér I, p. 455-458, 1992.

[17] YU.A. KUTOYANTS AND P. PILIBOSSIAN. *On minimum L_1 estimate of the parameter of the Ornstein-Uhlenbeck process.* Statistics and Probability Letters, 20, p. 117-123, 1994.

[18] C.PARR. *Minimum distance estimation : bibliography.* Commun. Statist. Theor. Meth.,A10(12), 1205-1224.

[19] T. PFAFF. *Quick consistency of quasi-maximum likelihood estimators.* The Annals of Statistics, v. 10, n. 3, p. 990 - 1005, 1982.

[20] V.N. LA RICCIA. *Asymptotic properties of weighted L^2 quantile distance estimators.* The Annals of Statistics, Vol. 10, $n°2$, p. 621-624, 1982.

[21] R.SH. LIPTSER AND A.N. SHIRYAYEV. *Theory of martingales.* Mathematics and its applications, Kluwer Academic Publishers, Netherlands, 1989.

[22] P.W. MILLAR. The minimax principle in asymptotic statistical theory. Lecture Notes in Mathematic, 976, p. 76-265, 1983.

[23] P.W. MILLAR. A general approach to the optimality of minimum distance estimators. Trans. Amer. Math. Soc., 286, p. 377-418, 1984.

[24] A.N. Shiryayev and V. Spokoinij. *Statistics of random processes.* (submited for publication)

[25] C. Sibeux *Convergence faible uniforme des semi-martingales vers un processus de diffusion.* Preprint , Department of mathemetics, University of Angers, n° 25,1996.

[26] L.Yu. Vostrikova. *On a weak convergence of likelihood ratio processes of general statistical parametric models.* Stochastics 23, p. 277-298, 1988.

[27] L.Yu. Vostrikova. *On weak convergence of parameter estimators of general statistical parametric models.* Séminaire de Probabilités, Rennes, p. 146-167, 1987.

[28] J. Wolfowitz *Minimum distance method.* Ann. Math. Statist., 28, p. 75-88, 1957.